



SDS PODCAST

EPISODE 941:

Multi-Agent Human

Societies, with Dr.

Vijoy Pandey



- Jon Krohn: 00:00:00 We have a fascinating future ahead of us wherein AI agents collaborate with us and each other to solve humanity's biggest challenges and improve quality of life for all. But a major hurdle toward this future is being able to trust agents. How are we going to do that? Welcome to the SuperDataScience Podcast. I'm your host, Jon Krohn. I'm joined today by Dr. Vijoy Pandey, a brilliantly thoughtful and well-spoken engineering researcher who heads Outshift a tech incubator inside of Cisco. In this episode, Dr. Pandey reveals an open source agent Agentic AI project that resolves the question of trusting agents and so much more. Enjoy this one.
- 00:00:38 This episode of SuperDataScience is made possible by Anthropic Dell, Intel Fabi, and Gurobi. Vijoy welcome to the SuperDataScience Podcast. It's a treat to have you on the show. How are you doing? Where are you calling it from?
- Vijoy Pandey: 00:00:54 I'm doing great. I'm calling it from San Jose, California. It's sunny outside and it's pleasant and it's a pleasure to be here.
- Jon Krohn: 00:01:01 The heart of Silicon Valley. That's fantastic. Yes. So Vijoy, you are senior vice president of something called Outshift by Cisco. Tell us about Outshift to give us some kind of, I think every listener is going to be aware of the Cisco brand, but maybe not the Outshift brand.
- Vijoy Pandey: 00:01:23 Yeah, so Outshift is an internal incubator within Cisco. So our entire mission is to incubate new products all the way from ideation to customer traction. So we're structured as startups with engineering, product marketing, customer support, sales, all roll into one, and we look at problem statements that are adjacent to Cisco's core businesses. So if you think about Cisco, like you said, we are familiar with the networking giant, the security, there's observability, especially after the Splunk



acquisition, and then there's collaboration with WebEx and so on. And so we look at personas and problem spaces that are adjacent to those four pillars, and that's where we incubate.

- Jon Krohn: 00:02:07 Fantastic. That was a great intro. I actually learned more about Outshift than I ever knew before. And within Outshift there's a particular initiative called AGNTCY that I think will be of interest to our listeners. And so this has a funny spelling. It's A-G-N-T-C-Y.
- Vijoy Pandey: 00:02:27 That is excellent. You get a hundred hundred.
- Jon Krohn: 00:02:32 I did that without looking at any resource that was from Memory AGNTCY. And yeah. So tell us about AGNTCY from Outshift by Cisco.
- Vijoy Pandey: 00:02:40 Yeah, I think before we even get into AGNTCY, the way this thing all came about was we were looking at a problem statement almost about two years ago where we said, agents are here, they're going to help us solve for all kinds of human work. And so how can we build a platform where agents and humans can collaborate with each other and solve for a business outcome, a social outcome, a physical outcome or a services outcome or whatever. And so that platform was missing and we strongly believed, I mean if you think about Cisco, we were one of the few companies that were behind the original internet and the way it transformed and it was open and interoperable and that's why it's successful. And so with this collaboration platform for agents and humans, we thought it should be like the original internet. And so we came up with this thesis called the Internet of Agents, and the openness comes from open source these days, standards, yes, but open source are the new standards. And so we said what better way to influence standard specifications and how it's deployed

and consumed done by taking internet of agents and making it open sourced. And that's what AGNTCY is.

- Jon Krohn: 00:04:07 Really cool. And before we started recording, you started to tell me about how you're excited for a future that will involve collaboration between teams of agents and humans. In fact, I guess you might not even think of it as teams of agents. It's just this multi-agent future where agents are available and collaborating with each other or with us on all manner of tasks. Vijoy, tell us about your vision for the future that's coming.
- Vijoy Pandey: 00:04:37 Yes, I'm a strong believer that in the future you'll have societies that are built off humans and agents, and these societies will help us solve for the biggest problems that face us. That is the understanding of the self, the understanding of the universe. So personally and selfishly, I'm looking for better medication. I'm looking for better materials. I'm looking to figure out how did the universe form, where are we headed? So that's where we are going towards, and us humans alone are not going to be sufficient in that journey. And so to solve for those bigger problems, we need these societies and teams that are multi-agent human societies. And these agents are not the narrow term that we use today, which is if you think about agents today, you think about agents in business software. I mean how boring when you think about agents and services and agents that write code for you.
- 00:05:40 I mean, that's great, but it's like how myopic and how boring. So you need to think about agents that are helping you discover materials and drugs and so on. Scientific discovery, you think about agents probably better than what Mark Zuckerberg about, but agents that help you socialize and interact with each other, agents that will help you of course solve a business and consumer needs and services, but most importantly



agents that are also embedded in physical form, whether you call it physical AI or embodied ai. And they help us accelerate all human work even in the physical domain. I mean, the first thing I want agents to do is do my dishes. I don't want my agents to paint a picture for me because that's what I enjoy doing. I want to play my guitars. I don't want spend time doing dish loading or washing my clothes. So I think that entire spectrum of discovery of work, physical work, of business work, social interaction, all will be solved through and accelerated through these multi-agent human societies. And that's where we all technologists need to aspire towards.

- Jon Krohn: 00:06:56 For people not watching the video version of this podcast, there are multiple guitars behind Vijoy. Are some of the things in frames hanging on the wall or any of those, your photos or pieces of art?
- Vijoy Pandey: 00:07:10 No, but I do dabble in photography as well. Although you see there are LPs and vinyl from the sixties. I'm a big, I spend money in audio, I spend money in photography. I spend money in a bunch of things that I want to continue spending on, and I want to develop agents that actually help me do the monotonous and the boring and the toil.
- Jon Krohn: 00:07:34 So you can have an agent do your podcast interview and you can be learning a new guitar solo.
- Vijoy Pandey: 00:07:39 I'll be in the background riffing and drinking some wine.
- Jon Krohn: 00:07:45 Nice. I like that. That is actually, that's something, that particular idea is something. One of my favorite guests that we've ever had on the show is a woman named Natalie Mom Bio, and she specializes in this idea of digital twins where you could theoretically have, I mean not theoretically, she actually, she works at bringing them to life today with a technology that we have today. So that you could, in a Slack channel or even in a video format,

have a digital version of yourself, be able to answer common questions, be able to maybe do a podcast interview for you. But yeah, so that's a really exciting vision for the future.

- Vijoy Pandey: 00:08:21 Can I interrupt you right there because...
- Jon Krohn: 00:08:23 Absolutely.
- Vijoy Pandey: 00:08:24 One of the things that we've been looking at is, so you said digital twins and digital twins can be twins of humans, they can be twins of systems, they can be twins of the world that we see around us, so physical twins, twins, twins of digital systems and humans. And this is an area of research. This is an area of how to solve for this agentic future that we just talked about that actually excites us the most because one thing that we've realized is that we are at a point in time where we have a set of tools that we were familiar with in the computing space that were deterministic in nature, and now we have a set of tools that are coming up that are probabilistic in nature. And this is using either traditional ML or energetic AI and generative AI and so forth.
- 00:09:18 So you have deterministic tools, you have probabilistic tools, and you can pick and choose. You can blend the words, and that's how you solve for problems. The issue is that the craze that is behind LLMs and energetic AI and generative AI is forcing people to just think about the istic tools and in some ways is ignoring the plethora of deterministic tools that really work well in a lot of situations. So the one area of research, and it's, it's not new, but you're seeing a resurgence behind that is neuro symbolic ai. And the whole notion behind that is just neural networks or the neuro part of it is not enough, just the symbolic or deterministic or knowledge graph or ontologies, whatever we might want to call it, the deterministic part of it is not enough, but the

combination thereof is where the magic happens. And so you talked about digital twins, and that triggers this thought where we actually came up with a whole bunch of use cases. One particular one in the network integration pipeline and validation pipeline that we built uses a digital twin of the network and then throws in a bunch of agents which are agent like generative AI and probabilistic in nature. And the combination of the two gives us speed and accuracy at the same time.

- Jon Krohn: 00:10:51 That does sound very cool. When you say symbols, it also makes me potentially think of symbolic learning, which could potentially vastly accelerate learning for machines and vastly reduce the amount of data that are required in order for machines to learn. And so this would be a bit more like how children learn where they don't need a million examples of cat versus dog images in order to be able to label the images correctly like a machine does. They're just able to see one or two examples and you correct them a couple of times. And so does that relate into this conversation?
- Vijoy Pandey: 00:11:29 Yes. So I think that is where the neuro symbolic movement started from, but I think you can expand that out by saying you can combine the deterministic worlds and the probabilistic worlds together, and then the outcomes that you get will have better accuracy. And of course the speed you expect to get also, it'll allow you to explore things in a better way than simple symbolic or deterministic systems as well.
- Jon Krohn: 00:12:02 That is all fascinating, but I want to get back to the idea of these multi-agent human collaborative societies a little bit. You mentioned how it won't just be boring business processes on the internet and the cloud on people's machines that are being handled. It sounds to me there you might also be talking about real world embodiments of agents like robots. Is that correct?



Vijoy Pandey: 00:12:25 That is correct. So the way we think about the way agents are going to come together and collaborate, and I'll give you one example, which is one of my favorite examples that actually came in through a pharma company and they were talking about this whole motion from wet labs to dry labs or silico discovery of drugs. And the way he was describing the process, this person was, they're using LLMs to first figure out a plan and the goal for what drugs they want to discover. So you start with an LLM, you go through some reasoning, some plan discovery planning, and you come with a plan. And so that's the LLM front end, so to speak, to this entire process. From there, you would go to a scientific foundation model, so something like a protein founding model where you an alpha fold where you explore all the possible scenarios for protein combination.

00:13:25 And that's the language that that foundation model is speaking and it's a pure exploration. And so based on that exploration, you tweak it, you prune it, you figure out what of those make sense to then go towards a wet lab. And so the wet lab is looking at robotics, embodied ai, taking those explorations and actually trying it out in real life in the physical world. And from there you would get into human trials and possibly animal trials as well. And underlying all of these are other agents that are looking for efficacy, they're looking for cost, they're looking for compliance. So all of the things that you need to do to make it a business and not just a scientific exploration. So if you think about that entire pipeline, you've got language models, you've got scientific foundation models, you've got embodied agents, you've got compliance and cost and safety, for lack of a better word, agents all coming together to solve for better drugs. And more and more of this is going to happen where is a platform that enables all of these agents to come together and collaborate. That's the problem that we went after was it's not siloed to one vendor. It's not siloed to one



cloud, it's not siloed to one vertical. It's really, really heterogeneous in nature and it's trying to bring all of those things together to solve for one big goal.

- Jon Krohn: 00:15:04 And so you're specifically talking about AGNTCY there,
- Vijoy Pandey: 00:15:06 So I'm talking about the eight of agents and how that translates to AGNTCY, which is open source manifestation of internet agents.
- Jon Krohn: 00:15:16 Gotcha. Gotcha, gotcha, gotcha. And so when we say internet of agents, that includes in the future these real world embodiments, it's just basically the internet of agents is the interconnectivity of all of these non-human intelligences that will be going around helping us on our screens or off our screens?
- Vijoy Pandey: 00:15:41 Yes. So when we think about the internet of agents is the embodiment of probabilistic agent software in all of these verticals that we just talked about. So whether it's social interaction, scientific discovery, B2B software services as well as physical embodiment in physical ai, robotics. And it's not just the communication aspect of it, but it's also how do we discover capabilities, how do we identify them, how do we compose them into workflows or in fact self forming teams? And then how do we help them communicate with each other? And then finally, how do we even have evaluate them because they're all here to do something for us. They're not just here to talk and communicate and have fun. So how do we evaluate them and make sure that they're going towards a common goal?
- Jon Krohn: 00:16:37 Right. I've got you. And so the internet of agents and then AGNTCY in particular provides an infrastructure layer for the AI era allowing all of this collaboration between agents on their own and between human agent collaboration?



- Vijoy Pandey: 00:16:56 Correct. And you can think about it in terms of you can deploy AGNTCY again as an open source manifestation of the internet of agents. You can deploy the components of AGNTCY within your organization to handle all the agents that are being developed inside your organization, or you can deploy it in such a way that agents from different vendors and different organizations can also come talk to each other, collaborate and solve for something. So it's like there's a microcosm it within an enterprise, but there's also a macro which lives outside a singular enterprise.
- Jon Krohn: 00:17:36 Cool. This is, so the AGNTCY framework, which is open source and which people will have a link to it in the show notes, again, it's spelled A-G-N-T-C-Y. That time I did look down and read it off a page. I wasn't know why I was so confident the first time, don't want to overdo my luck. So yeah, we'll have a link to that in the show notes. Open source. It sounds like I suspect you have an even bigger vision for what AGNTCY will be in the future, but already today it sounds like it blends together a lot of different kinds of functionality that different people have provided, like MCP provides tool use for agents. And then A two A is a framework from Google that is allowing inter agent collaboration. It sounds like with AGNTCY you are trying to provide an open source framework that allows all of these things together.
- Vijoy Pandey: 00:18:27 That's right. And I think, so if you think about how we took AGNTCY and moved it to the Lennox Foundation, the formative members for the AGNTCY project in the Lennox Foundation are of course Cisco, but there's Dell, there's Google, there's Red Hat and there's Oracle. Those are the formative members. And then there are about 80 other organizations, both small and large, that are contributing, using, and then deploying this at scale. So that's roughly where things are with AGNTCY. But also like you said, so we on the protocol side, we do collaborate with eight two



A as well as MCP. And we are a foundational member of the eight two A project that Google brought to the lf. So if you think about eight two A, we are a foundational member and Google is a foundational member of AGNTCY because we believe that fragmentation at this stage is really bad thing as it is people are having trouble deploying agents, building agents, getting some traction and outcomes out, especially in the enterprise.

00:19:35 And if you confuse everyone with multiple standards, multiple open source projects, we are not doing justice as technologists. So we all came together and made sure that when we say it's open and interoperable, we are truly being interoperable. And so if you think about the AGNTCY architecture, we actually show MCP as a protocol going from agents down to tool use and to data sources. And we show eight way as a protocol that sits between agents. Now we are Cisco, we are a networking company. We've done this a while, and we all know that protocols are protocols and there'll be many, many protocols. And so the way we think about AGNTCY is we love all protocols. We love MCP today and A two A today. There might be five others that might prop up. If you think about the internet, you have ip, you have TCP, you have UDP, even they were not sufficient. So the industry came up with quick. So there are many, many protocols that come along. I mean, there's HTTPS that sits above all of this. So there'll be many protocols and I believe that we are at the cusp of redefining that protocol stack that exists and we can dive in that direction if it makes sense.

Jon Krohn: 00:20:59 Yeah, let's do that a little bit. Terms like TCP IP for US data science listeners, we might not be familiar with that. HTTP, we have some idea because we see that in the web address. So you have some idea that it has to do with the way that webpages are sent over the web securely. But TCP ip, I mean having somebody from Cisco, you can probably explain that better than most people. And

critically, you have a PhD in computer science from uc, Davis that you got in the nineties. And so you have compared previously, you've compared today's agent ecosystem to the pre TCP IP days of the internet. So I'd love to hear first of all, TCP ip, I know it's really important for the way that the internet works, but I wouldn't be able to explain it. So I'd love it if you can explain it and then explain how the moment that we're today is analogous for agents.

- Vijoy Pandey: 00:22:03 So think about the early nineties. This was again when I was a bright-eyed undergraduate. Internet was just coming out. This was 1995 and we had dial up and I don't know if listeners here would remember it or even have experienced it.
- Jon Krohn: 00:22:26 I'm sure we will. I did.
- Vijoy Pandey: 00:22:28 And so if you remember that dial tone where you would connect the computer to the internet through a OL and a OL would come on cd, you would pop it in or maybe even a floppy and you would pop it in and it would dial into a phone line somewhere and you would get 56, and I'm going to mess this up, you get 56 kilo board or something like that. I don't even remember, but it used to be really slow, let's put it that way. Really, really slow internet connectivity from a single provider, A OL. And so you had these islands of connectivity and these islands never talked to each other. You had DECnet and Apple talk and I-B-M-S-N-A and a whole bunch of these other protocols on islands, and these networks actually did not talk to each other. It might seem surprising, but I'm on a Mac right now and I should be able to just talk to Mac if AppleTalk was the defacto king.
- 00:23:33 And if you're on a Windows machine, tough luck if you're on tough luck. And so TCP IP came in because surf in one of these meetings stood up and said, enough is enough.

You have B-S-D-T-C-P IP stack that was embedded in the BSD operating system from Berkeley, which was open sourced. And he came in and said, we are all going to standardize behind the TCP IP stack. And one of those books actually sits on my bookshelf over there. But that was a seminal moment in networking because, and Cisco was a big supporter of that. And what that did was it allowed packets from one box to flow to another box without it being siloed through a business need or a business requirement. So you truly open it up, you truly made the network interoperable. And that's how the internet was formed because everybody signed up behind T-C-P-I-P.

00:24:40 We are at a similar juncture for agent communication where yes, some of us are talking to each other, but some of us aren't. And like I said, we did not stop at TCP IP because it was invented in the nineties, I think late eighties, early nineties. The world has evolved. And so the requirements of TCP IP have changed over time. So the reason quick came about is we needed the same guarantees of communication that TCP provides. I need to ensure that you and I, when I say award does reach you and vice versa. So I need that guarantee. But then TCP is pretty heavy and I need that word to be transmitted to you, especially for communications very swiftly with low latency because the one thing people hate is delays when you see audio or video. And so something like Quick came about to marry the best of both worlds between TCP and UDP, so it gave the guarantees of delivery, but it also provided low latency interactive communication.

00:25:51 Similarly, what we are seeing on the agent side is yes, we have MCP, yes, we have a two A, but there are many, many needs for agent to agent communication. And over time we will see many protocols pop up. And as we create those protocols for a variety of needs, we need to make sure that there is a scaffold that exists around them, the

scaffolding of discovery of agents. This is like the DNS, you said data scientists only know the URL. So how does A URL translate to an IP address, which you don't really care about, but it does translate to an IP address. How does that happen? That happens through DNS. And so what is the DNS system for agents? How do we do agent discovery and capability discovery? That's a scaffold around T-C-P-I-P. That's a scaffold around A two A and MCP. So the discovery, the identity, of course, the messaging layer that we just talked about and the evaluation pieces that I mentioned earlier, this is the scaffold that needs to exist regardless of the number of protocols that we all as an industry come up with.

Jon Krohn: 00:27:05 Cool. I understand all of that now, and I understand the analogy. It makes a lot of sense to me. Something that I understand is a big issue for agents today in this early ecosystem, this early internet of agents where we don't have all the protocols sorted out. One of the issues that I understand we run into is identity and access management. So a IM is the abbreviation used for that often, which is kind of a fun thing like I am. And you're providing your identity. You're saying who you are. So I am, that's a clever thing. I'd never noticed that before. And so traditional I am internet or identity and access management systems often break down when applied to AI agents. Why is that?

Vijoy Pandey: 00:27:52 So identity and access management, by the way, I didn't even think of it as I am, which is a great analogy there. But the reason the breakdown is so far all IM services and systems have worked with humans or deterministic software and agents are a combination of the two. So agents are like humans in the way they communicate, in the way that they are using NLP, but also in the way that they act and behave and change their persona. They're all probabilistic in nature. You cannot predict, I mean you ask Chad GPT the same question five times over, you'll

get five different answers. They're very, very stochastic in nature in terms of both intent and in terms of outcome, but they also operate at machine speed and scale. So unlike vOy, of course is hallucinating half the time who unpredictable and is actually speaking natural language, but thankfully vOy is not running at machine speed and scale.

00:29:04 So if you give vOy sub access control to do some damage, you only give that access control for a short period of time. So you can think hours, let's say, or even days, great, but agents are acting on behalf of vOy today might act on the behalf of Jon in the next minute. And they're doing this at machine speed and scale. And how do you build an identity and access management system to handle this probabilistic behavior at scale? And that is the big problem. So typically the positions that we've used in IM are three kinds of positions, role-based, access control, attribute based, access control and relationship-based access control. These are the things that we've used in Im to identify what access to give a human or a piece of software and role-based access control works again, because humans have certain roles and they don't deviate from those roles in a certain context within our organization. Agents like we just discussed are not in the same time boxed, role boxed environment. And so they will differ and we need to treat them differently.

Jon Krohn: 00:30:28 Alright, so okay, now I understand why Im systems break down for AI agents and it totally makes sense to me there that a big part of the issue is that it's the machine scale, it's the machine speed that you can, if you provide permission to agents, you can all of a sudden have so much activity happening, so much network activity, so much real world impact in a way that a single human never could. So just it scales up the security risks

associated with identity and access management. So yeah, what is the solution?

Vijoy Pandey: 00:31:05 So identity and access management for agents needs to go back to basic first principles and think about what is it that we are truly trying to achieve by giving somebody access to a protected piece of software protected data object or whatever it is that you're trying to give access to. So the basic principles behind that is looking at the task that you're trying to achieve, the tool that you're trying to access or the transaction that is taking place. So we defined this whole notion of T back. So T stands for task tool, transaction based access control. So that's the T back. Sometimes we want to call it T three back because there are three Ts in this equation. But this is going back to first principles and saying instead of trusting an agent by saying this is an agent for v Oy, so let me trust that this agent is going to behave like vjo oy for the next five days.

00:32:15 Instead of doing that, I'm going to go back to the first principles of zero trust and say vi's agent, what is it that you're really trying to achieve? Which task are you trying to do? And based on that task, I'll figure out what level of authorization and access you need. I'll give you a token to go and do that task at that level of authorization and then I quickly bring you down before the next task or transactional tool access takes place. So it's going really, really granular figuring out what a task could be elevating your privileges if needed for that task or tool access or transaction. And then quickly bringing you back because it's not just the speed, like you said Jon, but it's also the probabilistic nature of the agent because it's not deterministic software, it's machine, it's software and it's the data being fed into it and the probabilistic outcomes and the probabilistic communication that is taking place with the agent that makes it behave differently from one second to the other.

Jon Krohn: 00:33:26 Right. Got you. Yeah, the probabilistic nature of the machines makes it particularly difficult to trust machines because even if you actually trust the person who's provided this capability, probabilistic things, the probabilistic nature of LLMs means that you could get unexpected outcomes anyway, which could lead to security issues. So you have this zero trust AGNTCY, ZTA framework, and you have these two different types of access control. There's tac or as you said, T three back because it's task tool and transaction based access control then, so that is what you were just describing, where based on some specific task or some specific tool or just one particular transaction that you need an agent to be involved with tac, you're just providing control for that particular task, that particular tool, that particular transaction. And so that differs from the other way that you could be providing access control, which is role-based. And it's that role-based ROLE, that role-based access that is, I guess that's traditionally what we were used to. You used to have Vijoy would've access or Jon would have access or whoever would have access. That was role-based control. But now in this AGNTCY world, it makes more sense to have this tac which is much more specific.

Vijoy Pandey: 00:34:58 Yes. And I think the way to think about this is you still will have role-based access control for humans. You'll still have attribute based access control and relationship based access control for software and humans and services that are deterministic in nature. But for agents, we want to move towards T three back or tba task tool, transaction based access control. And we need to bridge these two worlds because the old world is not going away. Jon will have role-based access control, VJA will have access control, which is role-based, but my agent and your agent, and sometimes it's the same piece of software that is taking one persona at some point in time and some other persona at some other point in time. And so



we do need to bridge the RB systems of the world to the TAC systems of the world for this multi-agent human societies to come together from the identity standpoint.

00:35:56 I mean if you're just looking at that narrow piece of the vision. And the other thing is right now everything that, I mean there's a translation that's happening here. So I think you mentioned sometime in the past, but we all know this, where humans are going to be part of the loop for a long time to come. So even though there'll be agents in a workflow, in a self forming team, whichever way you can think, tasks will evolve or teams will evolve. Humans will be part of the loop or human in the loop is going to be a thing for a long time to come. And so that's the other reason why role-based and task tool, transaction-based access control need to talk to each other because there'll be points in this workflow where you will need to punt to a human and say, are we doing the right thing? Are the agents doing the right thing? Do you approve of this? And so that's the other reason why we need to bridge these two world together. But as far as agents are concerned, we have to move towards this first principle based access control, which is tac.

Jon Krohn: 00:37:07 Nice. That makes a lot of sense. When we're working with something like tac, it sounds like it might be tricky. I can't wrap my head around exactly how permissions are granted. Just in time when you need to grant permission to an agent to be able to do a particular task and then you need to revoke that afterward. That sounds like it could be complicated. How do you handle it?

Vijoy Pandey: 00:37:32 We have to bring in a whole bunch of infrastructure around this notion of TAC to enable the end goal, which is the zero trust for agents, which is I give you permissions to do something specific for that just in time. Just in time. And then for that duration of that task two or transaction and then I revoke that token or revoke

those permissions the moment you're done. So what else do you need? So in my head, the equation runs like this, which is zero trust AGNTCY is I don't trust any agent and I just trust it once it's proven that it's supposed to do X, Y, or Z for the duration of that X, Y or Z, and then I revoke those permissions. So that is zero Trust for agents is a combination of the availability of trust, a task tool, transaction based access control.

00:38:29 So all identity providers, authorization servers need to support tac. That's step one. We need to have a passing entity for all communication that's taking place between agents and agents and humans that can pass that communication, that pass that discourse and figure out the tasks or the tool access or the transactions that are taking place between agents. So that's the second piece because that'll help us define what those tasks tool access and transactions are. So there's a semantic passing element and then your adjusted time comment basically implies that you get a token, you do that task in a very contained sandbox jailed environment and then you're taken out the moment you're done. So the analogy I draw here is I want to access a safe, which has a lot of money, but I want to withdraw \$10 from the safe. Now you can give me since I vjo oy, you can give me access as vjo oy to go and open that safe.

00:39:48 But then you can say, you know what? There are other people's money in that safe, so I'm going to give Vijoy just enough to withdraw cash for the next 10 seconds and then move out. So that's like a task-based access control. But then I need to pass the communication that's happening between myself and somebody else where we are talking about withdrawing 10 bucks and say, okay, Vijoy is allowed only to withdraw 10 bucks and that is the task he's doing. So let me just give you access for that \$10 withdrawal and not sit around to withdraw all of the money from the safe. So that's the task based passing



that needs to happen. And finally, I'll let you in into the sandbox environment, give you that authorization to withdraw \$10 and then I'm going to shut the door. I don't trust you beyond that point, I will not let you linger around. That's safe. So that is a sandbox runtime environment that needs to happen. So is the hooks in identity providers to provide task-based access control is a semantic passing of the discourse of the communication to figure out what that task is and then a runtime sandbox environment to just do that task with that authority and then get out. So those are three things that need to come together for zero transfer agents to have

Jon Krohn: 00:41:13 Semantic parsing, ephemeral runtimes and human in the loop approvals. And overall, you gave me a really clear picture now of what this all involves. One thing that I guess I'm still, that I still don't quite get logistically, you said that you won't trust the agent even for the particular task tool or single transaction that you're going to approve it for until the agent has proven itself. How do agents prove themselves trustworthy in the first place?

Vijoy Pandey: 00:41:43 This is where the entire pipeline comes to the picture. So we are looking at identity, which is the first stumbling block that everybody's running into. And so one of the things that we're seeing is in the AGNTCY framework, the identity piece is the problem to solve first even before you can start deploying agents at scale within the enterprise. But then as you pointed out, there are other aspects of trust, there are other aspects of semantic parsing. So there are these other aspects of their entire pipeline that we need to solve for. So coming back to trust, the simplest way you can start with is saying, is there a directory somewhere that allows me to discover agents that are trusted? So I want to find a financial agent not from a particular vendor maybe, but also if there are 10 vendors, I want the best of breed financial agent from a vendor that is highly reputable and highly trusted.

00:42:51 So the simplest version of the question, the answer to your question is, is there a directory which tells me which agents are trusted agents? But then the next question is, so that's where the directory comes in. So we have a directory where you can discover agents through capabilities and through things like reputation and trust. But then the next question would be how is that trust enforced or attributed? Is it crowdsourced? Crowdsourced could be one thing. So 10 people have used it. It's like the Apple app store which says five stars trust or trust pilot score. And it's like, yes, it's awesome, but if you want to be a little bit more mathematical and provide some rigor, then you go to the last pillar of that four pillar thing that I talked about earlier, discovery, identity, communication and evaluation. You look towards evaluations and you look towards evaluating agents and multi-agent workflows and saying, over time, I've built trust by evaluating this agent. And that trust can then feed back into the directory's reputation score and say, yep, all good to go till something perturbs in the system because these things are constantly evolving.

Jon Krohn: 00:44:11 That's cool. I get it. So this is kind of like when I'm on Amazon and there's some product that I want to buy, and one vendor on Amazon is providing the product for \$5 cheaper than another, but the one that's doing it for \$5 cheaper as a 70% approval rating from people who have made purchases before and the other one has a hundred percent, I'm going to pay an extra \$5.

Vijoy Pandey: 00:44:33 And in that example, you and I who are buying this product are the evaluators, and Amazon is the discovery directory engine, which has the trust reputation score attached to it because we as evaluators that are actually feeding back into Amazon and saying, this thing sucked, so bring it down even further or this is awesome.



Jon Krohn: 00:44:54 So does all of this that you've now covered in the episode, so things like the zero trust AGNTCY framework, things like tac, does that allow me as a data scientist or as an AI engineer now to maybe be able to get more of my agentic projects approved by enterprise security teams?

Vijoy Pandey: 00:45:13 True. And that is the whole goal behind this, right? So when we thought about the internet of agents and AGNTCY in the open source, the notion behind the entire project was to ensure that agentic workflows and self forming agent teams and getting agents to do business outcomes or perform business outcomes, the bar to that is lowered or becomes easier. And that's where we started. So identity is a big stumbling block and a problem to solve. Evaluation is a problem to solve. If you think about an enterprise like Cisco, we are dealing with agents from a Salesforce or from Cisco, Microsoft, ServiceNow, all of these things need to come together to build a simple sales funnel, for example. And so these agents are all sitting in their own clouds, different vendors, different clouds, different intent and outcomes. So if somebody says, give me the best X, what does best mean when some agent comes back and says, here's the best X with 90% confidence, is it 90% well understood across all agents or is it vendor by vendor thing?

00:46:26 And so looking at discovery identity, bringing them together and helping them collaborate and evaluating them, these are the four steps of the pipeline that we saw enterprises struggle with on a day-to-day basis. So to your question, yes, data scientists, I'm an engineer, we are all focused in on the outcomes that we want to deliver on our product, on the thing that we are working on. But if you take a step back and to your point, think about adoption, there are these really hard brownfield heterogeneity problems that we need to solve and AGNTCY is out there to solve for it in an open source, open interoperable way.



- Jon Krohn: 00:47:11 I like it because for me, I like to be able to focus on capabilities and I don't want to have to worry too much about security myself. And so I'm glad when folks like you, an AGNTCY come along and you solve my security problems for me in a, yeah, you just solve all the problems for me and I don't have to handle them. It's great. Turnkey is the term I was looking for there.
- Vijoy Pandey: 00:47:36 So it is an open source project and one of the ways open source gets adopted and grows and is successful is if it is not a biased viewpoint. And so the way we think about AGNTCY and open source in particular is we want practitioners, we want developers, we want vendors, we want operators, we want consumers and customers, all personas and roles to come in, play with AGNTCY, figure out what's working, what's not, whether it's helping, whether it's not is a documentation up to snuff any which way possible, help us grow that community, help us utilize the projects, provide real feedback. So even though you would expect and we want to provide a turnkey solution through the product stream that Cisco is building based on AGNTCY, the open core part of AGNTCY, the open source part of AGNTCY could benefit a lot from developers like you, practitioners like you and the listeners here to come in and just contribute in whichever way that you deem necessary.
- Jon Krohn: 00:48:49 Nice. And so if people want to get started today with AGNTCY, they like the sound of what you've been describing. They want to be able to have a more interoperable system for their agents to work with each other, to work with humans, and for this to work in a way that happens across whatever kind of framework they're using. It works with crew ai, it works with land graph, llama, index, whatever. So obviously we're going to have the GitHub repo in the show notes for people to go to. Is there anything they should know? I'm seeing for example, that there's something called coffee AGNTCY, which is a



fictitious coffee company that helps developers, scientists understand how components in the AGNTCY ecosystem can work together. Maybe is that a place that you'd recommend people starting or where should people get started with AGNTCY if they're curious?

- Vijoy Pandey: 00:49:41 You can absolutely go to AGNTCY.org and I'll spell it out again this time. So it's A-G-N-T-C-Y.org and that's where you will get access to documentation, you'll get appointed to the git repo and you can join and have fun. But like you pointed out, coffee AGNTCY is an open source reference application that we built and it's truly a coffee AGNTCY. So we've got multiple coffee suppliers. It's a supply chain problem that we show. There are coffee manufacturers scattered across the globe. There's a full supply chain and then there's a coffee shop that's trying to actually sell coffee. And we picked this because A, we love coffee and B, this is a pretty complicated supply chain example, which brings in all of the complexity that any enterprise goes through. And it's a reference application because it's a real application, but the entire code is open sourced so you can plug and play various components within the application.
- 00:50:41 So it's not just the AGNTCY components within coffee AGNTCY, of course it size eight to a, it has NCP, but you can plug in Cassandra and see how that works. You can plug in, I don't know, Tipco as a messaging bus and see whether TIPCO will work as an inter agent collaboration platform. Or you can plug in various IDPs like an Okta or a duo. And so this is one place where you can bring in your environment into this reference application and then swap things in and out. Maybe it's your code, maybe it's somebody else's code, and see the benefits of using these various components and build out a real world example before deploying it into production in your environment.



- Jon Krohn: 00:51:31 Nice. Very cool. Thank you for doing this for us. Why are you doing this for us? Why does Cisco invest and all these other vendors, you mentioned tons that are involved in this AGNTCY project. What's in it for Cisco in the end?
- Vijoy Pandey: 00:51:48 Few things. First of all, we truly believe in the open interoperable future of multi-agent human societies. I mean, at least that's my belief. And Cisco is a big proponent of an open interoperable internet and has always been, but that's a vision statement, that's a belief statement. There are business ramifications to it as well. And the business ramifications come again from the belief that open systems provide the maximal value for every participant in the ecosystem. So whether you are a vendor like us, whether you're an operator, whether you're a developer, whether you're a customer or consumer, every persona in that ecosystem benefits if that ecosystem is open from the dollar perspective, customers will get cheaper. Interoperable products, vendors will make more money because this thing is widely adopted. So if you think about the value prop for every business, they will get maximal value if the ecosystem is open and interoperable. So that's the reason why we're in it. And also we feel that building this it of agents is pretty much our birthright and participating in this is something that we do well because we understand distributor systems distributed computing at scale, and the next step towards this word computing is actually distributed computing. And so we want to go after that future and we can build it in the right way.
- Jon Krohn: 00:53:33 I love that makes a lot of sense. Cisco are the masters of the internet. Why not be the masters of the internet of agents that it's coming? Makes a lot of sense. Before I let you go, actually, I was about to start getting into the final questions that I ask all my guests, but right before that, one final one popped into my head, Vijay, what's next for this initiative? What's next for AGNTCY?



Vijoy Pandey: 00:53:55 AGNTCY so far? I mean we touched upon this a little bit in the earlier conversations where we are building the scaffold for these different agents sitting in different organizations with different roles and personas on different clouds from different vendors to all come together to solve for a business or a consumer or a scientific or physical work outcome. The problem is that right now, whatever exists out there in AGNTCY and in other protocol layers and other frameworks like H two A and MCP and other frameworks that you mentioned, we are dealing with this heterogeneity, this complexity, this inter collaboration at a certain layer. So we are dealing with what I call a syntactic layer of complexity. So this is like saying agent one and agent two, you and I speak German, you speak Japanese, we are trying to figure out the structure of our language and we just want to standardize on English.

00:55:07 So we say, okay, noun before verb or verb before noun, there's our sentences constructed. And that translates to a framework. So whether you built on a Ang chain or Ang for you, whether you built on a crew AI or whether you're built on a bedrock or the agent SDK from Google, these are frameworks and these are just syntactical representations and you need to make these things interoperate. But there's a bigger problem at hand. And the bigger problem, especially if we are true to our vision, which we are, we believe in this vision of multi-agent human societies, the bigger problem is to understand each other semantically. So even if though we are speaking English, do I really, are we standardized on what we mean when we say certain things? So when we say best, when we say 90%, when we use a certain terminology or a phrase, are we standardized on that and can we understand the meaning behind the communication that's happening?



00:56:12 So the semantic layer is actually the next set of problems to go after and when you go after the semantic layer, that's why I said there are many, many more protocols that will appear, but it's not just the communication part. The moment you start getting with the semantic layer, you're dealing with knowledge, you're dealing with cognition as the problems get really, really interesting when you take knowledge and cognition and you start spreading that around and you start thinking of it in terms of societies. So that's what's next when it comes to 800 agents per se.

Jon Krohn: 00:56:47 Nice. Getting closer and closer to that vision that you mapped out at the beginning where agents and humans are working together to solve the biggest problems facing society and making the world's better, better place to live in. I certainly share that techno optimistic vision. Some people may think we're too optimistic, Vijay, but I don't know, you can't stop the tech. It's going to keep coming and there are certainly downsides to technology, but by and large, if I had to pick a time to be living in history, I want to be born now not a hundred years ago or anytime before that.

Vijoy Pandey: 00:57:23 As a technologist. I mean Jon, I mean yes, every piece of technology has a downside or a negative aspect, but as technologists, I believe that we should build technology to solve for the problems that technology creates. We need to solve it through fundamental principles and make things better. And that's just good business as well,

Jon Krohn: 00:57:44 For sure. Alright, so I already kind of alluded that this was going to happen. I snuck one last question in there. Every guest I ask them before I let them leave the show for a book recommendation, what do you have for us?

Vijoy Pandey: 00:57:58 My favorite book is The Life of Pi.



- Jon Krohn: 00:58:01 Oh yeah, life of Pi. Cool.
- Vijoy Pandey: 00:58:02 And there is a big reason for that because I believe we live in a very probabilistic world. I believe that just like quantum physics, and that's the other thing that we dabble in day in, day out. Within Outshift we are looking at the quantum internet and quantum networking. I believe that there is a superposition of states and we are just measuring and looking at one state and that's what we are in right now. And so you can believe in either science or you can believe in faith or a combination thereof. It's up to you. It's up to all of us. But whichever way you are leaning, there is a story that you believe in. There's a scientific story that you believe in. There's a story around faith you believe in or somewhere in between. That's a story. And the life of Pie to me was the eyeopener in the sense that we all believe in stories. It's also like Sapiens. I mean we are a big believer in stories, but the Life of Pi told me that believe in the story, which is more fantastical, which drives more excitement and which will make you live a more interesting life because you can always decide to believe in the boring story, which is what fun is there in that. So if you're going to believe in some story, believe in one that is going to drive excitement for you and just stick to it.
- Jon Krohn: 00:59:35 That was a cool book pitch. I love Sapiens and it sounds like I'm going to now have to check out Life of Pi. So we're blending deterministic and probabilistic models to get the best outcomes from our enterprises. We can blend science and faith to get the most interesting outcomes for us as individuals. Cool. And yeah, very last thing, Vijay, is how should people follow you for your brilliant thoughts after this episode? Obviously we know to go to the AGNTCY website and the AGNTCY GitHub repo, which I'll have in the show notes, but what about you personally or are there any other links from Shift or something that our listeners should be following?



Vijoy Pandey: 01:00:13 So I think AGNTCY.org works for AGNTCY. You can go to Outshift.com. That's to follow everything that we're doing in internet of ages, but also the quantum internet side of the house, which is even more fascinating and mind blowing. But to follow me personally, of course I'm on LinkedIn. I'm the most prolific on LinkedIn, so you can follow me on LinkedIn, which is slash in slash Vijoy on Twitter on X. And then I have a website which is a little bit dated, but I'm trying to update that. But I think LinkedIn is the best place to follow.

Jon Krohn: 01:00:48 Nice. Thanks for Vijoy. I agree that LinkedIn is the place where you can probably find most of our listeners and most of our guests these days. It's pretty interesting how that's happened. Great to have you on the show. I really enjoyed this episode. You are brilliant and fun to speak to. Thank you so much for taking the time out of what is surely a very busy day for you.

Vijoy Pandey: 01:01:08 This was a fascinating conversation. Thank you, Jon.

Jon Krohn: 01:01:13 In today's episode, Dr. Vijoy Pandey covered his vision for multi-agent human societies where agents and humans collaborate to solve everything from scientific discovery to physical tasks, freeing humans to focus on creative work. He talked about AGNTCY Cisco's open source platform for the internet of agents that enables agents from different vendors and clouds to interoperate and collaborate. He talked about the importance of blending deterministic and probabilistic tools rather than abandoning proven deterministic approaches for pure AI solutions. How digital twins of humans systems in the physical world are key to building reliable agentic systems that can be tested safely before deployment. And the evolution from syntactic interoperability between agents to the harder problem of semantic understanding where agents comprehend the meaning behind their communications. As always, you can get the show notes including the



transcript for this episode, the video recording, any materials mentioned on the show, the URLs for Vijoy's social media profiles, as well as my own at superdatascience.com/941.

- 01:02:16 Thanks everyone on the SuperDataScience podcast team are podcast manager, Sonja Brajovic, media editor, Mario Pombo, partnerships manager, Natalie Ziajski, researcher Serg Masís, writer Dr. Zara Karschay, and our founder Kirill Eremenko.
- 01:02:31 Thanks to all of them for producing another exceptional episode for us today for enabling that super team to create this free podcast for you. We're so grateful to our sponsors. If you are ever interested in sponsoring the show, you can find out how to do that at jonkrohn.com/podcast. Otherwise, share, review, subscribe, but most importantly, just keep on tuning in. I'm so grateful to have you listening and I hope I can continue to make episodes you love for years and years to come. Until next time, keep on rocking it out there and I'm looking forward to enjoying another round of the SuperDataScience Podcast with you very soon.